

Probabilistic Rail Vehicle Localization with Eddy Current Sensors in Topological Maps

Stefan Hensel, *Student Member, IEEE*, Carsten Hasberg, *Student Member, IEEE*,
and Christoph Stiller, *Senior Member, IEEE*

Abstract—The precise localization of rail vehicles is fundamental for the development and employment of more efficient train control systems in security, logistics and disposition applications. Current research in train navigation systems tries to solve the task with an increasing number of onboard sensors or additional infrastructure installations in combination with satellite navigation (GNSS). Both approaches are cost intensive and rely on undisturbed satellite signals, commonly not given in railroad applications. In contrast, we describe a novel, single sensor, onboard localization system in this contribution, based on a newly developed eddy current sensor (ECS). We outline an onboard localization system within a probabilistic framework, with special attention on signal processing for speed estimation and pattern recognition. In particular, we employ Bayesian methods such as hidden Markov models for turnout detection and classification and, in a final step, sequential Monte Carlo sampling to combine the extracted information in a topological map to obtain a reliable position estimate.

Index Terms—Train Localization, Eddy Current Sensor, Hidden Markov Models (HMM), Sequential Monte Carlo (SMC).

I. INTRODUCTION

TRAIN localization systems nowadays operate with a high amount of track side infrastructure installations. Future train operating systems, e.g. the European Train Control System (ETCS) on level three [1], are based on an onboard localization of rail vehicles to enable more efficient disposition techniques such as *moving block*. Therefore, current research focuses on reliable onboard train localization systems. Inspired by localization techniques in air and automotive traffic, the commonly chosen system architecture combines onboard speed measurement with global navigation satellite systems (GNSS), commonly GPS, to estimate a reliable position information [2]. One drawback of this approach is the relatively low accuracy of inductive odometric measurements and the GNSS reliability, given that tracks contain tunnels and roofed station halls and are often situated either in urban areas, forests or valleys with dense vegetation and high buildings. Hence, the reliability of the obtained solution is insufficient

for the requested security and technical standards. Current disposition methods such as ETCS on level two try to improve accuracy with additional infrastructure installations to gain recalibration points, which leads to high track maintenance costs.

Eddy current sensors (ECS) are commonly used for the measurement of inhomogeneities within ferromagnetic materials (e.g. see [3]). Based on this principle, we developed a sensor system capable of precise non-contact and slipless speed measurement in rail vehicles [4], [5]. Besides the increase in odometric measurements the sensor additionally allows the detection and classification of turnouts. In contrast to vision based systems for track and turnout recognition [6], the ECS working principle is robust against weather influences and day time. The sensor outperforms accelerometer based systems as [7], [8] with the possibility to detect turnouts passed unbending and to classify individual turnouts.

Rail networks naturally possess a large amount of turnouts, e.g. the German rail network has a total length of approx. 38000 km containing about 72000 turnouts. We propose to use these as beacons that are intrinsically distributed according to the need of local precision, i.e. lots of turnouts within stations that contain a lot of parallel tracks for a landmark based localization.

The sensor information is one-dimensional and therefore suited to be employed with topological maps. These are, in contrast to precise metric maps, easy to obtain or can easily be created by any person that knows the track. Localization with topological maps is commonly researched in indoor robotics and examples can be found in [9] or [10]. We illustrate how a combination of robust velocity estimation and reliable turnout detection are sufficient for track specific onboard train localization with ECS rendering other relative and absolute positioning systems obsolete.

For this purpose we consider and setup the complete localization system, which can roughly be subdivided into three parts. Speed estimation, the recognition of turnouts and the final information fusion within a map. To cope with the requirements of the system such as heavy duty conditions of rail vehicles, different vehicle types, independent noise sources, very large rail networks, and changes in rail infrastructure, we employ probabilistic methods. The advantages of a stochastic concept are shown in many robotics [11], tracking [12] and machine learning [13] applications. The Bayesian framework we employ for turnout extraction and map based localization is intuitive, robust, and easy augmentable to incorporate additional sensors. With this our system can be employed as a standalone solution on side tracks or in combination with other

Manuscript received March 1, 2010; revised August 26, 2010, April 13, 2011, and June 22, 2011; Date of publication July 29, 2011. This work was supported in part by the German Federal Ministry of Economics and Technology. The Associate Editor for this paper was M. Brackstone.

S. Hensel, C. Hasberg and C. Stiller are with the Institute of Measurement and Control Systems, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany, e-mail: {carsten.hasberg, stefan.hensel, christoph.stiller}@kit.edu.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TITS.2011.2161291

onboard sensor systems for safety applications in ETCS level three. Nonetheless, in this contribution we try to emphasize the individual capability of our localization system.

The remainder of the paper is organized as follows: Section II will introduce the eddy current sensor system and its application to speed estimation. We propose a novel method that alleviates acceleration influences to significantly improve the velocity estimate quality based on cross-correlation of the signals, especially when driving in low speed. Recognition of turnouts with hidden Markov models is subject of Section III. It is shown how probabilistic models can be employed to detect and continuously extract turnouts. Afterwards, the classification of the detected samples is analyzed with respect to signal preprocessing, parameter adaptation and model selection. Velocity and turnout information must be combined in a most optimal way within the map, which is subject of Section IV. We resolve ambiguous position hypotheses, that are represented by an *a posteriori* probability density function, with sequential Monte Carlo (SMC) methods. This allows handling uncertainties in track association that occur while passing turnouts or due to misclassifications. The final Section gives a summary and conclusions of the localization concept as well as possible augmentations.

II. EDDY CURRENT SENSOR SYSTEM (ECS)

A. Sensor setup

Eddy current sensors are commonly used to detect inhomogeneities in the magnetic resistance of ferromagnetic materials (e.g. see [3]). This basic approach has been further developed and adapted to possible applications on railway vehicles. These include speed measurement and pattern recognition. For correlation of two signals the eddy current sensor system consists of two identical sensor devices, each built up with a transceiver coil and two pick up coils. Both sensors are placed sequentially within a housing, mounted approximately 10 cm above the rail head. Figure 1(a) displays the principle of a single sensor unit: The transmitter coil E excites a magnetic field H_E , that induces eddy currents in metallic materials like the rail. The eddy currents induce an antipode magnetic field H_{EC} , that generates the voltage $u_{P1}(t)$ and $u_{P2}(t)$ within the pick-up coils P1 and P2 respectively. By interconnecting them differentially, the output signal $s(t) = u_{P1}(t) - u_{P2}(t)$ is a measure for rail inhomogeneities. These mainly result from rail clamps, turnouts and other irregularities, e.g. cracks or signal cables (for details see [4]). The signals $s_1(t)$ and $s_2(t)$ represent a stochastic process. Clamps produce a stationary process for rail vehicles driving on open tracks with constant velocity. Turnouts, cables, and metallic clutter represent non-stationary signal components, whereas both parts are superimposed by a noise process that can be regarded as zero mean white Gaussian noise. The overall signal comprises a high signal-to-noise ratio (SNR), given that preprocessing low pass filters are installed in the sensor hardware [14].

B. Velocity estimation

The described working principle is, in contrast to vision based systems or doppler radar sensors, widely unsusceptible

to environmental perturbations and, because of the differential setup, robust against systematic influences. These properties are highly desirable for a reliable speed measurement under rough railway conditions. Velocity estimation can commonly be achieved via cross-correlation of the two sensor signals $s_1(t)$ and $s_2(t)$, that are idealized depicted in Figure 1(b). First approaches, intended and optimized for hardware realization, apply a closed loop correlator (CLC) assuming a known sensor distance l and a measured time difference Δt (for details see [4]) and [15].

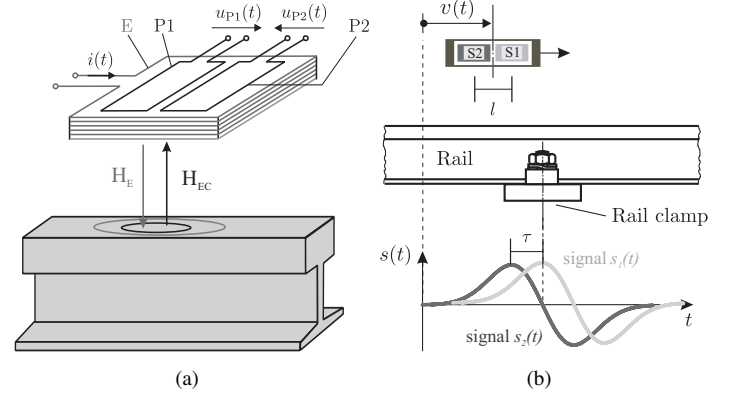


Fig. 1. (a) Single ECS sensor S_1 (b) Example signal of ECS (two sensors) $s(t)$ when crossing a rail clamp.

The presented approaches rely on the assumption of a stationary stochastic process, which holds for constant velocity within the cross-correlation interval. Whereas this assumption is correct in most situations, it is heavily violated in low speed manoeuvres, where large changes in the relative velocity may occur. This is unfortunately the case in areas of interest, i.e. within stations, where most turnouts are present. The need of a precise spatial signal $s(x)$ for pattern recognition (see Section III) makes it necessary to apply a velocity estimation, that can cope with these situations. Therefore, we augment the common cross correlation analysis by a correlation based estimation of the current train acceleration. With this, the signal can be resampled to eliminate correlation window averaging effects on an accurate velocity estimate.

The basic idea is to find the time shift Δt of the signals $s_1(t)$ and $s_2(t)$, which corresponds to the maximization of the cross correlation according to

$$\Delta t = \arg \max_{\tau} (E \{s_1(t - \tau) \cdot s_2(t)\}). \quad (1)$$

We assume a constant acceleration \hat{a} in a signal of equal length before and after resampling. With an average velocity v_0 within the interval T_M , the traveled distance d_{T_M} becomes

$$d_{T_M} = \int_0^{T_M} v(t) dt = \int_0^{T_M} [v_0 + \hat{a}t] dt = v_0 T_M + \frac{1}{2} \hat{a} T_M^2, \quad (2)$$

and using signal length equality

$$d_{T_M} = v_m N T_s = \tilde{d}_{T_M} = v_0 N T_s + \frac{1}{2} \hat{a} (N T_s)^2, \quad (3)$$

for a discrete sequence of length N and sampling time T_s .

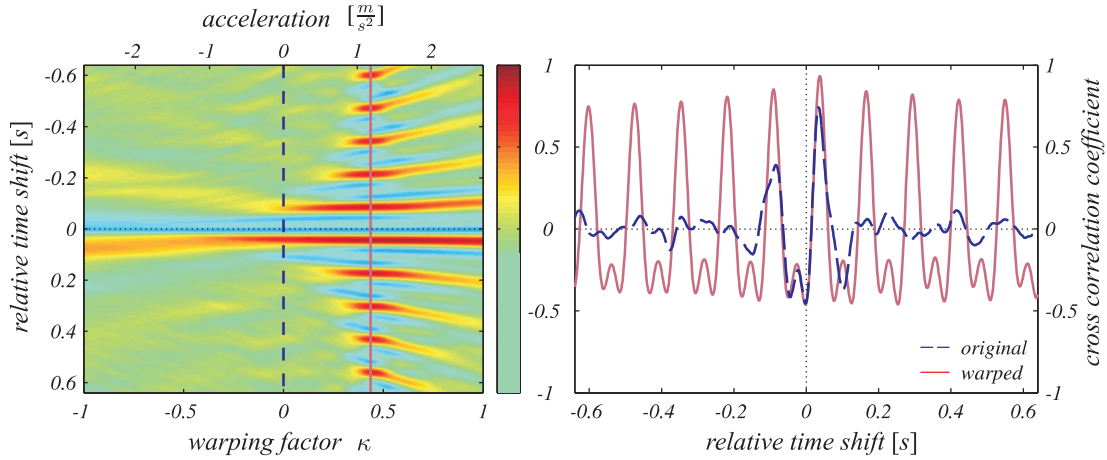


Fig. 2. Cross correlation coefficients plotted against warping factor κ using a 4s window. By varying κ , one can determine κ_{max} that locally maximizes correlation coefficients. The right figure displays two example coefficients as functions of κ (indicated by the corresponding lines in the left figure).

The association of an original sample n_o to the resampled output sample n_i is expressed with (3) and the auxiliary variable

$$\zeta := \frac{v_0}{\hat{a} \cdot T_s}, \quad \zeta \in (-\infty, -N] \cup [0, \infty), \quad (4)$$

which results in

$$n_i = -\zeta + \zeta \sqrt{1 + \frac{2n_o}{\zeta} + \frac{N \cdot n_o}{\zeta^2}}. \quad (5)$$

A more practicable domain is obtained by reparametrization of (5) with a *warping factor* κ . Under consideration of extreme values ([16]) it follows

$$\kappa := \frac{4 \Delta n}{N} = \frac{N}{2\zeta + N}, \quad \kappa \in [-1, 1]. \quad (6)$$

This gives κ a limited range of values with $\kappa \in [-1, 1]$. For the determination of \hat{a} and v_0 , ζ is replaced in (6) with

$$\zeta = \frac{1 - \kappa}{2 \kappa} N. \quad (7)$$

Thus, the signal resampling for the purpose of maximizing the cross correlation coefficient is characterized by κ , that describes the signal straining. Figure 2 shows the influence of κ for the cross correlation quality, and its applicability for acceleration estimates. After resampling, the correlation maximum is determined with parabolic approximation for subsample accuracy. The average speed $\hat{v}(t)$ and acceleration $\hat{a}(t)$ within a given interval is calculated with

$$\hat{v}(t) = \frac{l}{\Delta t}, \quad \hat{a}(t) = \frac{2\kappa \cdot \hat{v}(t)}{L \cdot T_s}, \quad (8)$$

where L is the number of samples within the correlation integral and T_s the sampling time. The estimates of $\hat{v}(t)$ are subsequently interpreted as observations and merged in a Kalman Filter (see [17], [12]) with constant acceleration model to track the final velocity estimate $v_{est}(t)$. In contrary to a CLC this approach allows speed estimates down to zero velocity. The additional knowledge of the Kalman Filter acceleration estimate $a_{est}(t)$ is used for validation of the calculated $\hat{a}(t)$.

Further details on error propagation and results for the velocity estimation are outlined in [16].

The high accuracy of this method allows for subsequent velocity integration to determine the covered distance in sufficient quality, which is first pre-requisite to solve any localization problem and calculation of spatial signals $s(x)$, with $x = \int_0^T v_{est}(t) dt$ as covered distance in time T .

III. PATTERN RECOGNITION WITH HMM'S

First step for enabling pattern recognition is the transformation of the time signal $s(t)$ into a spatial signal $s(x)$ to yield invariance against velocity influences, that strain the turnout signal when passing over several times. Afterwards, we apply a model based approach, that assumes a common physical turnout structure to model the various kinds of turnouts [18]. Figure 3 depicts the chosen basic modeling type and a corresponding ECS signal $s(x)$ of a real turnout. In our

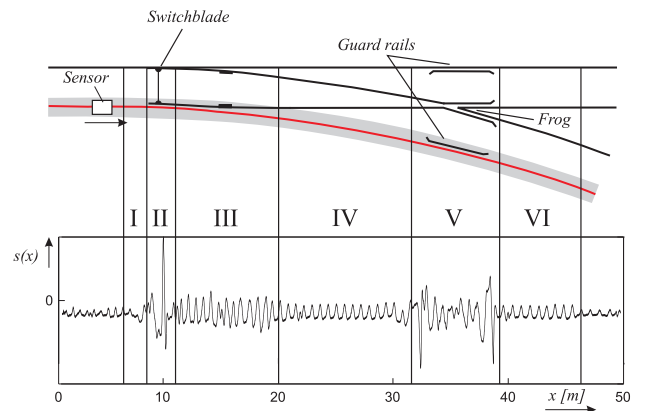


Fig. 3. Turnout model depicting standard turnout, straightened signal $s(x)$ (recorded with an ECS mounted above the right railhead), and the six assumed segments of a turnout.

contribution we assume that any turnout can be segmented into six areas, each characterized by a specific length l_{seg} and average amplitude \bar{a}_{seg} . In Figure 3, an example for passing over is depicted where the train bends off to the right. An

ECS sensor system mounted on the right train side yields the shown signal $s(x)$. In contrast to velocity estimation, turnout detection can be conducted from the information of a single channel. For convenience we will consider only one signal $s(x)$, whereas implicitly both channels can be used separately for the sake of redundancy.

The signal is segmented as follows: The starting area I , where commonly welding points are situated, the switchblade actuator in segment II , the following third segment with the bending area of the switchblade, an interconnecting segment IV , in which rail clamps are laid that show higher amplitude compared to open track clamps depending on the sleepers. The, due to its discriminative character, important segment V where either frog or, like in the given case, a rail guard induces the signal shape, followed by a final sixth segment in which again turnout rail clamps characterize the signal. A noteworthy result of this model is, that each possible turnout driving direction, in the following called facing right/left and failing right/left, provides a different and unique sequence of the main turnout parts, switchblade, frog and rail guard (see Table I). The detection of turnouts makes use of the characteristic features of each component sequence and tries to allocate it in the signal. Assuming a stationary Gaussian stochastic process [19] for open track signals, turnouts give need for a non-stationary signal interpretation. The non-stationary signal is separable into sub processes, e.g. for a turnout six, whose realizations are connected by the stationary clamp signal.

Therefore, the recognition is solved in a probabilistic way. Detection and Classification are formulated according to Bayes rule as

$$P(\mathcal{T}_m|\mathcal{O}) = \frac{P(\mathcal{O}|\mathcal{T}_m)P(\mathcal{T}_m)}{P(\mathcal{O})}, \quad (9)$$

where $P(\mathcal{T}_m|\mathcal{O})$ represents the Bayesian *a posteriori* probability of recognizing \mathcal{T}_m out of a set of $\mathbb{T} = \{\mathcal{T}_1, \dots, \mathcal{T}_M\}$ possible classes, given the feature vector sequence $\mathcal{O} = (O_1, \dots, O_L)$ of length L . The class set \mathbb{T} for detection includes rail clamps, basic turnout types and possible disturbances that are commonly present on rail tracks. Classification aims to determine a particular turnout \mathcal{T}_m out of a set of $4 \cdot M$ classes that represent all possible passings for an overall of M turnouts.

The true turnout conditioned probability density $P(\mathcal{O}|\mathcal{T}_m)$ in (9) is not known. We model this density with $P(\mathcal{O}|\lambda_m)$, where the parameter set λ_m represents a hidden Markov model (HMM). The modeling and estimation of λ_m is subject of the following sections. In the following we introduce the basic model descriptions and their notation for clarification of subsequent sections.

A. Hidden Markov Models

HMMs are stochastic models widely used in speech processing [20], bioinformatics [21], time series analysis [22] and other machine learning and pattern recognition fields. In this contribution we propose to model the sequence of turnout segments as a Markov chain. With this, HMMs that represent a doubly stochastic process are predestined to be employed

for detection, given their capabilities in coping with both, variations in length and amplitude within a given signal.

We adopt the notation of Rabiner [23], to describe a HMM as a two staged stochastic process. A Markov chain of N possible states and initial state distribution vector π , obeying $\sum_{i=1}^N \pi_i = 1$, is completely defined by its state transition probability matrix $\mathbf{A} = \{a_{ij}\}_{N \times N}$, with $\sum_j a_{ij} = 1$. a_{ij} defines the transition probability $P(q_t|q_1 \dots q_{t-1}) = P(q_t = j|q_{t-1} = i)$ for any state q_t at time step t . A second process generates symbols of a given set at every time step, of which only the emitted series of symbols is visible depending on the states taken at every time step. This probability $P(O_t|q_t)$ is defined in the *emission matrix* $\mathbf{B} = \{b_{jk}\}_{N \times K}$, where $b_{jk} = b_j(v_k) = P(O_t = v_k|q_t = j)$ and $\sum_j b_{jk} = 1$. For turnout segmentation and classification, HMMs with continuous probability density functions according to $P(O_t = \mathbf{x}|q_t = j) = b_j(\mathbf{x})$ with $b_j(\cdot)$ obeying $\int_{\mathbf{x}} b_j(\mathbf{x})d\mathbf{x} = 1$ are used. A HMM is hence completely determined by its parameter set $\lambda = (\pi, \mathbf{A}, \mathbf{B})$. For further details on HMMs and there applications we refer to [20], [24] or [21].

B. Detection of Turnouts with HMMs

The ideal turnout model described in the beginning of Section III involves several unknown parameters and additive noise. The length l_{seg} depends on different turnout types (for examples see [18]), speed estimation inaccuracies and additional individual structural features for each turnout. The average amplitude \bar{a}_{seg} is liable to filter effects, unpredicted bogie movements, clamp types, and varying sensor settings. Additional noise sources involve thermal noise and measurement effects. With this, we assume, due to the central limit theorem [19], an additive zero mean white Gaussian noise with individual variance for the true parameters l_{seg} and \bar{a}_{seg} . As depicted in Figure 4, the transformation in the two dimensional feature space discriminates each segment sufficiently by its mean and variance. Mean and variance of the depicted Gaussians are won from recorded test data via maximum likelihood (ML) estimation [13] and the Gaussian assumption successfully confirmed by a χ^2 -test with significance level of 5%. The biggest discriminative power lies in the main component models II and V that are also depicted in Figure 3. They show a good separability to each other and can be distinguished from open track areas and disturbances such as cables, metallic clutter or unique infrastructure such as bridges. The latter can have similar shape to turnout components, e.g. switchblades, but can be distinguished by knowledge of the turnout set-up. Turnouts are thus represented by both, a distinct sequence of components and their respective features. HMMs are perfectly suited to model the underlying turnout sub-component sequence and their conditioned features. In addition, if the signal component space is discrete and finite they are optimally in a Bayesian sense.

1) *Signal preprocessing for turnout detection:* In a pre-processing step the time domain signal is transformed to spatial space. This alleviates the velocity dependent signal straining proportional to the quality of the speed estimation (see Section II-B). Afterwards, exploiting the signal property

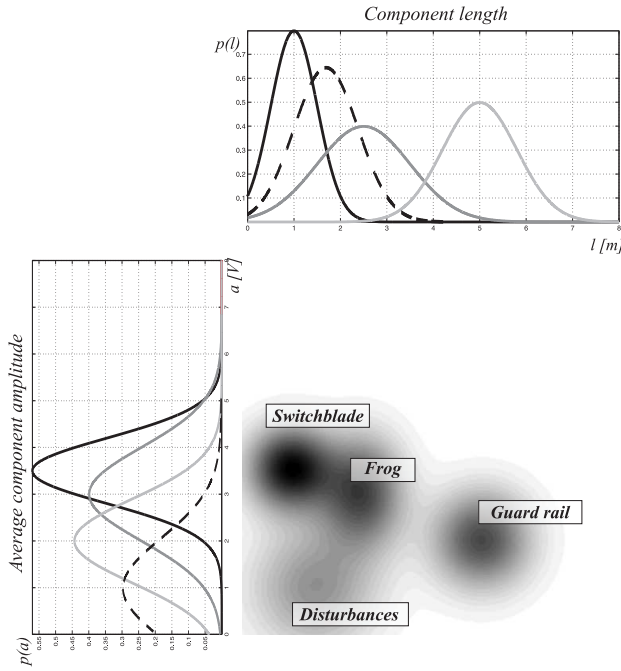


Fig. 4. Feature space with Gaussian assumption for l_{seg} and \bar{a}_{seg} , displaying main turnout components (solid lines) and common disturbances (dotted line).

of having zero mean, the signal power $P_s(x)$ over an interval of length $2I$ is calculated via the variance, determined as

$$\overline{\sigma^2}(x_0) = \frac{1}{2I} \int_{x_0-I}^{x_0+I} s(x)^2 dx = P_s(x_0). \quad (10)$$

Result is a non-periodic signal that serves as observation sequence \mathcal{O} for the HMM.

2) *Feature representation in detection HMMs*: After pre-processing the signal, the HMM framework from Section III-A must be adapted to detect turnouts according to the assumed physical model. The signal $s(x)$ represents a discrete stochastic process of infinite length that is comparable to continuous speech data or gesture recognition [25]. The HMM topology best suited to process sequential data of this kind is a linear left-right model, depicted in Figure 3. This topology assumes only self transitions a_{ii} or transitions to a state higher than the current ($a_{ij} = 0 \forall j < i \wedge j > i - 1$). Subsequently we represent each segment in Figure 3 as one state of the HMM as depicted in the upper half of Figure 5.

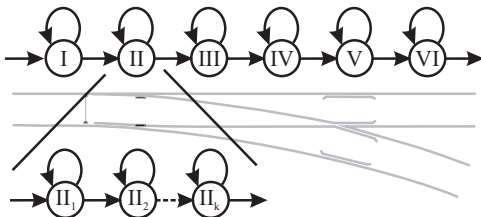


Fig. 5. Applied left to right model structure with *state tied* substates.

The detection HMM must represent the two features length and average amplitude. Assuming independence, we separately model \bar{a}_{seg} in the emission matrix and l_{seg} in the transition matrix. This keeps \mathcal{O} univariate and formalizes the transition

density calculation. However, a common drawback of the basic HMM modeling is the geometrically distributed state duration model that must be improved to represent l_{seg} adequately. In our approach, the expectation of state duration corresponds to component length and must be adapted accordingly. We employ the *state tying* technique [20] to approximate the Gaussian distributed l_{seg} . This implies replacing each state $q_t = i$ with an amount of k sub states, each sharing the same emission density $b_i(\cdot)$. Henceforward we can statistically interpret the duration as sum of k independent geometrically distributed random variables. The resulting probability mass function for this kind of state duration is the negative binomial distribution defined as

$$P(l_{seg} = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}, \quad (11)$$

where $P(l_{seg} = n)$ is the probability mass function over staying n time steps in a state until the first time k successes occur within a sequence. The quantity p is the probability of success in a Bernoulli experiment and can be calculated via the state transition probability a_{ii} , given state $q_t = i$. By interpreting the number of necessary successes as sub states k , the expectation of 11 becomes

$$E_{negbin}\{l_{seg}\} = k \cdot (1 - E_{geo}\{l_{seg}\}), \quad (12)$$

where E_{geo} is the expectation of a single sub state geometric probability. The interpretation of k as design parameter allows to place the mode and shape of the resulting distribution as close as possible to the ones in our Gaussian model assumption. By doing this, the complete information of the segment length l_{seg} is coded within the transition matrix \mathbf{A} .

The distribution of the average signal amplitude \bar{a}_{seg} is coded within the emission matrix and is expressed with continuous state emission models as mentioned in Section III-A. The parameters of each Gaussian $b_j(\cdot)$ are estimated via *maximum likelihood* estimation of training samples. These can be recorded in a laboratory test bed to avoid test drives.

3) *Online state estimation*: As mentioned at the beginning of the Section, a turnout can be driven over in four directions, each one comprises of a characteristic sequence of turnout parts. The possible sequences of turnout parts are listed in Table I. Each turnout passing direction is associated with a

TABLE I
SEQUENCE OF MAIN TURNOUT PARTS FOR DIFFERENT DRIVING DIRECTIONS

Driving direction	1 st Component	2 nd Component
facing right	Switchblade	Guard rail
facing left	Switchblade	Frog
failing right	Frog	Switchblade
failing left	Guard rail	Switchblade

sub-model $\lambda_{1..4}$. These HMMs represent all different turnout types and their individual characteristics and distinguish only in direction which is sufficient for a correct turnout start determination. This information is used for localization and in the subsequent classification step as *a priori* information. Assuming rail track characteristics such as double clamps on

wooden sleepers, bridges, metallic clutter and electric installations, additional failure models are necessary. The detection model operates with two failure models $\lambda_{5,6}$ that are used to detect longer areas with mediocre average amplitude and shorter disturbances with higher amplitudes (examples are depicted in Figure 7). Figure 6 shows the final HMM structure with six sub-models interconnected by a sleeper model GS . After defining the topology, the sequence \mathcal{T} of the most

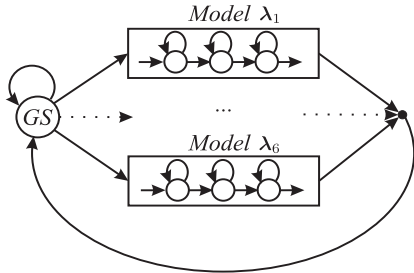


Fig. 6. HMM continuous recognition model. Six submodels $\lambda_{1..6}$ are connected via a *glue state* GS that models open track areas.

probable events given the set of events $\mathcal{T} = (\mathcal{T}_1, \dots, \mathcal{T}_c)$ is estimated. The solution space \mathbb{T}^* incorporates the set of all possible sequences given \mathcal{T} leading to

$$\hat{\mathcal{T}} = \arg \max_{\mathcal{T} \in \mathbb{T}^*} P(\mathcal{T}|\mathcal{O}), \quad (13)$$

where the posterior $P(\mathcal{T}|\mathcal{O})$ corresponds to

$$P(\mathcal{T}|\mathcal{O}) = \frac{P(\mathcal{O}|\mathcal{T})P(\mathcal{T})}{P(\mathcal{O})}. \quad (14)$$

The likelihood $P(\mathcal{O}|\mathcal{T})$ is again approximated by $P(\mathcal{O}|\lambda)$, the prior can be used to code syntactic relations for event sequences. With the *min-sum* or Viterbi algorithm [26] the optimal path in the sense of a *maximum a posteriori* (MAP) estimation subject to the minimum sequence error probability [24] is calculated continuously. In contrary to MAP estimator with minimum symbol error probability (e.g. the *sum-product algorithm*) only reachable paths are allowed, which is crucial for modeling the event sequence within turnout sub-states. The observation sequence \mathcal{O} must be restricted to remain computable. The algorithm adjusts the length of \mathcal{O} to start and end on open track and therefore adaptively to the detection of possible events within it. The justification is depicted in Figure 7. The correct result for the given observation sequence is shown in the upper half. Two subsequent turnouts, one crossed facing right the other facing left are situated behind a road crossing detected as disturbances. The correct main turnout component sequence is *switchblade* \rightarrow *guard rail* \rightarrow *switchblade* \rightarrow *frog*. In the lower half possible errors with a too short observation sequence window are depicted: While the window fully covers the first turnout (depicted in the lower middle cut out), the turnout is correctly detected. If the sequence shifts forward, the cut out in the lower right represents \mathcal{O} that starts at the end of the first turnout and ends at the beginning of the second turnout. Although this results in a valid combination of turnout components, namely guard rail and switchblade, a not existing turnout passed failing left is detected.

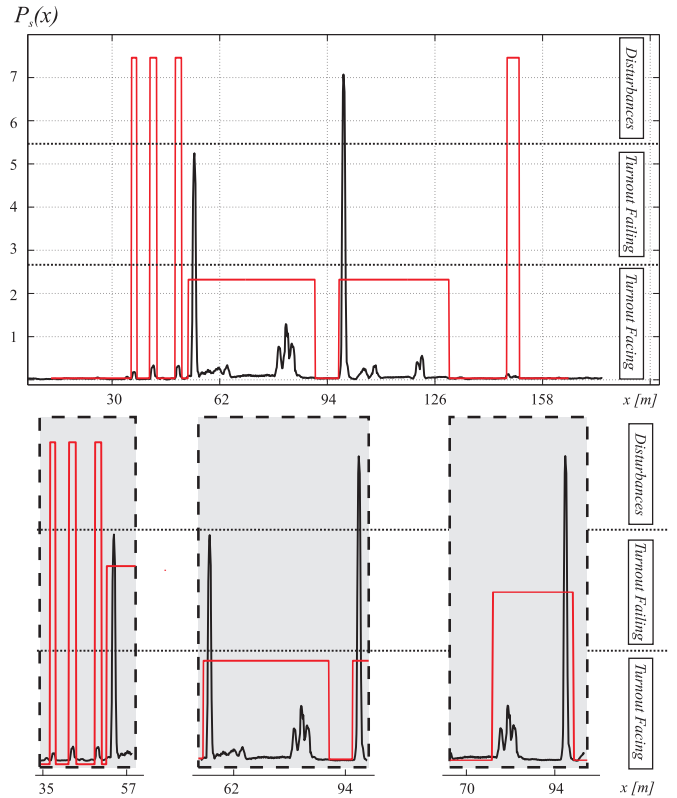


Fig. 7. Example segmentation results. The lower sequences clarify the problem of sufficient Viterbi path length and the importance of component sequence coding in HMMs. In the lower left the beginning turnout is false classified as facing, in the lower middle one can see that the component association can lead to sensible but yet wrong results. The lower right shows the same cutout as the middle one with a larger window and so a longer Viterbi path which results in a correct classification.

To prevent the appearance of these *phantom* turnouts, the length of \mathcal{O} is adapted until the sequence starts and ends on open tracks. The observation sequence now corresponds to a sliding window of minimum 100 meters on open tracks, where in stations with many turnouts the size can be up to 600 meters. Within the window, the forward probabilities (see (21)) are computed incrementally, while the Viterbi path is updated in every step for the whole sequence. The detection algorithm performance is in real time up to a sequence length of 800 meters, calculated with MATLAB on a 2 GHz laptop processor with 3 GByte RAM.

C. Classification of Turnouts with HMMs

The first HMM stage described in the previous subsection yields segmented sequences corresponding to possible turnouts. Although this is already sufficient to localize specific positions within a track plan, the ECS signals can be used for further classification. The common rail infrastructure makes it possible to reduce possible turnouts into basic types. But, especially on side tracks a lot of individual features such as slightly different placed signal boxes, cables, or other distinctions, exhibit additional discriminative features. We use this attribute to classify individual turnouts and the passing over direction in a given track system. Again, we apply HMMs

to cope with variations in the signals from bogie movements, sensor deterioration, and velocity estimation errors.

1) *Signal preprocessing for turnout classification*: In contrary to the preprocessing described in Section III-B1, where individual turnout attributes are smoothed to be allocated into basic types, a different signal preprocessing is chosen for classification to better distinguish individual features. With regard the non-stationarity nature of the turnout signals three different signal processing techniques were implemented. A cross validation comparison [27] of spatial signals, power signatures, *short time Fourier transformed (STFT)* signals and the application of the *discrete Wavelet transform (DWT)* [28], the superiority of the latter was identified. This coincides with its intrinsic capability to cope with local non-stationarity. An example of the wavelet transformation of a real turnout is depicted in Figure 8. Three scales from the transformed signal are taken as feature vector and henceforth employed as observation sequence for the classification HMMs.

2) *Classification and model adaptation*: Output of the preprocessing is a feature vector for every detected turnout that represents an observation sequence \mathcal{O}_j for the second HMM stage. Given J passings for a turnout one can assign \mathcal{O}_J sequences to a HMM Λ_m , with $m = 1 \dots M$ and $M = 4 \cdot \mathcal{M}$ for \mathcal{M} turnouts. In this notation each driving direction of a turnout is represented by a single HMM. The association of the sequences is manually done for initial test drives. Henceforth each newly classified sequence can be used for the parameter estimation, what allows an adaption on slowly changing influences, such as ECS deterioration or track abrasion. The implementation is realized with the Baum-Welch algorithm [29] adapted for multiple training sequences of each class. The last K classified sequences with associated to each turnout are used to estimate the model parameters Λ_m . This ensures a constant computational effort and a sufficient repression of misclassified samples within the training set. While the detection HMM has only a single Gaussian as emission density, the classification emission densities \mathbf{B}_m are designed to express individual features more accurate. Under consideration of the constrained training data for rail vehicle applications, we employ *semi-continuous hidden Markov models (SCHMMs)* [30]. These are best suited if only few training samples are available and the emissions may be modeled as mixture of multivariate Gaussians with a shared amount of basis functions.

The initialization for the iterative Baum-Welch training is made individually for transition probabilities, model size, and emission densities. The initial values for the transition matrices \mathbf{A}_m are chosen according to the self transition probability and the average length \bar{T} of the training set. Is the state transition treated as success and \bar{T} as number of trials, the probability for being in a state k in a left-right topology is described with the binomial distribution according to

$$P(X = k) = \binom{\bar{T}}{k} a_{ii}^{\bar{T}-k} (1 - a_{ii})^k. \quad (15)$$

With respect to expectation and variance of the distribution, we approximate the initial state number N for each model

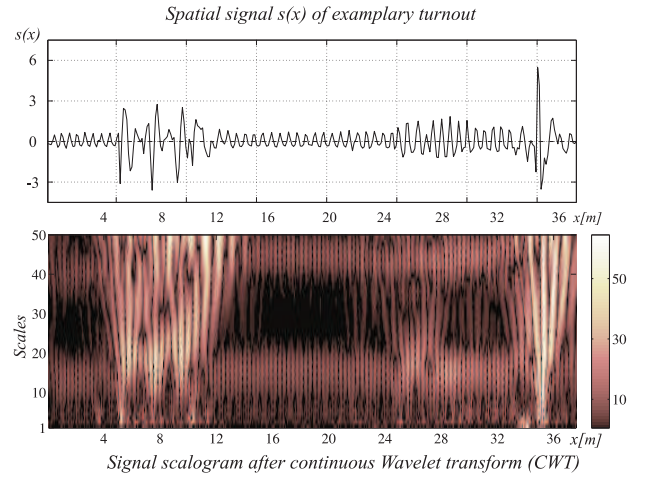


Fig. 8. Spatial signal $s(x)$ and corresponding scalogram of a turnout.

with

$$N = E(X) + 2 \text{Var}(X) = \bar{T} (1 - a_{ii}) + 2\bar{T} a_{ii} (1 - a_{ii}). \quad (16)$$

The model size is hence a function only dependent of the initial self transition a_{ii} . This ensures that all free parameters are used by the designed models to adapt the training samples. In addition, a balanced weighting of emission and transition probabilities is ensured. When compared to analytical model selection techniques, (16) estimates the number of initial states comparable to the *Akaike information criterion (AIC)* [31] used with the additional penalty term of Sugiura [32] (AIC_C), according to

$$AIC_C = \underbrace{-2 \ln \mathcal{L}_{\max}(\Lambda | \mathcal{O})}_{AIC} + 2N + \frac{2N(N+1)}{\bar{T} - N - 1}, \quad (17)$$

where $\mathcal{L}_{\max}(\Lambda | \mathcal{O})$ is the average of the maximum likelihood values of all training, test, and validation sequences for the model given its parameters and the number of states N with

$$\mathcal{L}_{\max}(\Lambda | \mathcal{O}) = \frac{1}{J} \sum_{j=1}^J p_{\max}(\mathcal{O}_j | \Lambda). \quad (18)$$

The *Bayesian information criterion (BIC)* [33] underestimates the needed number of states for the given application too strong, resulting in a lower classification performance. Equation (16) allows for a good model size estimation without calculating the *log-likelihood* as necessary in the mentioned analytical model selection criteria. This substantially reduces the computational complexity that is impractical even for very small rail networks with one hundred turnouts comprising an average length of $\bar{T}=400$ samples. Our estimate assures a sufficient model length to cover all features, prevents an over fitting, and is easily applicable for new turnout models of any size.

For the initialization of the continuous Gaussian mixtures in \mathbf{B} we apply the *Linde-Buzo-Gray (LBG)* algorithm [34].

After estimation of Λ_m , the classification is conducted by choosing the model that maximizes

$$\mathcal{T}^* = \arg \max_m \{P(\Lambda_m | \mathcal{O})\}. \quad (19)$$

Here, \mathcal{T}^* represents the associated turnout and its driving direction. Assuming a uniform prior over all possible turnouts, the posterior $P(\Lambda_m|\mathcal{O})$ is proportional to the likelihood and calculated with the forward algorithm [23] according to

$$p(\mathcal{O}|\Lambda_m) = \sum_{j=1}^N \alpha_T(j), \quad (20)$$

$$\alpha_t(j) = \left(\sum_{i=1}^N \alpha_{t-1}(i) \cdot a_{ij} \right) \cdot b_j(O_t), \quad (21)$$

where $\alpha_t(j)$ represents the probability of being in state $q_t = j$ at time step t , for likelihood computation.

IV. LOCALIZATION IN TOPOLOGICAL MAPS

Rail vehicles are not steerable and can be modeled with one degree of freedom preset by the rail network. With this, the train position is determined by the current track segment, the position upon it and the current driving direction. The ECS solves this with the combination of the velocity estimate $v_{\text{est}}(t)$ described in Section II and turnout extraction described in Section III. Yet, the information must be merged within a map. We employ topological maps as natural choice to represent a rail network. In our contribution a static map with known distances is assumed. The map can be set up with basic information of edge length and turnout connections that is accessible for every rail network. Although we imply that the on-board localization is solely based on the ECS system, our probabilistic framework can easily be enhanced with additional sensors, e.g. satellite navigation systems or inertial measurement systems to fulfill further security aspects.

A. Map representation

Topological maps are graph-based abstract representation of the environment. They are an adequate choice for the given application due to their intuitive understanding, the scalability, and compact representation [9]. We interpret the turnouts as vertices V and the connecting rail tracks as edges E in a graph. Figure 9 displays a map commonly used in signaling centers. It contains information of the distances between turnouts and is enhanced with track specific features, such as road crossings or platform positions. This kind of map is easily transformed into a directed graph, represented by the adjacency matrix G . In addition, if the turnouts are known in advance, each turnout \mathcal{T}_s with $s = 1 \dots M$ is associated with four HMMs Λ_m , where for convenience $m = s, i$ with $i = 1..4$. The graph deduction and turnout association is depicted in detail in Figure 10. The

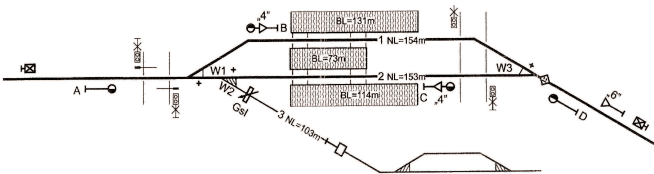


Fig. 9. Topological representation of a railway station on a side track.

nodes V and edges E in G are augmented with additional

information, such as turnout coordinates, metric distances, or distinct features for classifiers, and their relative position in the graph.

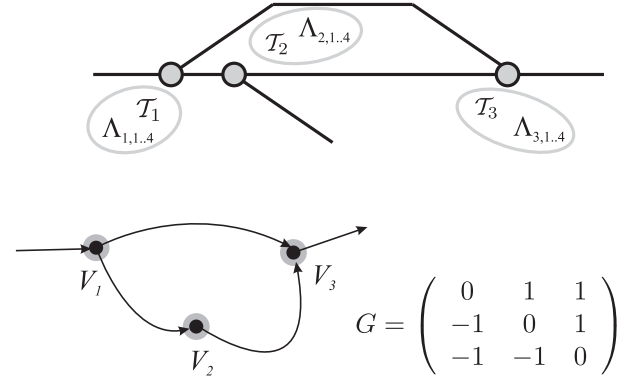


Fig. 10. Association of HMMs to specific turnouts and transformation of given map from Figure 9 into graph representation.

B. Information Fusion and Tracking

Positioning in topological maps corresponds to a one dimensional problem with multiple hypotheses conditioned on turnout bending decisions. This results in a multimodal position probability. Common object tracking within a Kalman filter framework [12] can only handle one hypothesis and is not applicable to this task. We employ *sequential Monte Carlo methods (SMC)* to recursively update a multimodal probability density function (pdf) over all possible train positions within the considered rail network. SMC approximate the conditional density

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t) \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})d\mathbf{x}_{t-1}}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}. \quad (22)$$

This pdf formulates the recursive Bayes filter of the state \mathbf{x} at time step t given all measurements \mathbf{y} up to the time t . Given a set of P samples or particles, in each discrete time step t associated with its weights w_t^p obeying $\sum_p w_t^p = 1$, one can approximate (22) with

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_{p=1}^P w_t^p \delta(\mathbf{x}_t - \mathbf{x}_t^p). \quad (23)$$

In this contribution we apply the sequential importance resampling (SIR) algorithm, for further details on SMC methods we refer to [35] or [11]. The linear model is expressed in state space form according to [36] as

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B_u u_t + B_f f_t \quad (24)$$

$$\mathbf{y}_{t+1} = h(\mathbf{x}_{t+1}, e_{t+1}), \quad (25)$$

with the current relative position on a graph edge x_{rel} as state variable x_t . The velocity estimate according to Section II-B is used as system input $u_t = v_{\text{est}}(t)$, where $B_u = B_f = T_s$ with sampling time T_s . The uncertainty f_t is defined as

$$f_t = e_t^{\text{est}} + e^{\text{mo}}, \quad (26)$$

and contains velocity uncertainty $e_t^{\text{est}} \sim \mathcal{N}(0, (\sigma_t^{\text{est}})^2)$, estimated by the Kalman filter and additional constant Gaussian white noise $e^{\text{mo}} \sim \mathcal{N}(0, (\sigma^{\text{mo}})^2)$ for model inaccuracies.

Each particle is additionally augmented with two *pseudo states*: The track segment ID_t , corresponding to an edge E within G , and its current driving direction dir defined as binary variable. The pseudo states explicitly determine the position in the map depending on x_{rel} .

The measurement model (25) is nonlinear and reflects discrete measurement events such as turnout detections. We assume a constant zero mean Gaussian distribution for measurement uncertainty $e_{t+1} \sim \mathcal{N}(0, \sigma^2)$, reflecting variations in the detection HMM stage or misclassification. With this, (25) becomes

$$p(y_{t+1}|x_{t+1}) = \mathcal{N}(d(x_{t+1})|\mu = 0, \sigma^2). \quad (27)$$

The term $d(x_{t+1})$ in (27) expresses the distance to the nearest node, given the current state and edge length $l(ID_{t+1})$ according to

$$d(x_{t+1}) = \begin{cases} l(ID_{t+1}) \cdot x_{t+1}, & x_{t+1} < 0.5 \\ |l(ID_{t+1}) \cdot (x_{t+1} - 1)|, & x_{t+1} > 0.5 \end{cases}. \quad (28)$$

The state space formulation combined with SMC approximation allows an elegant incorporation of turnout driving restrictions stored in the map G . Detection and classification probabilities are directly integrated in the particle weights in (23). The complete formulation is kept compact and can be augmented easily.

For localization it is sufficient to subsequently detect turnouts and estimate the driven distance between to exclude step by step false hypotheses. For this case only velocity estimation and the detection HMM (see Section III-B) are necessary to accomplish the localization. If turnouts are additionally associated to individual HMMs as described in III-C, an instantaneous localization is possible. The employment of the classification HMMs comes at the cost of needed prior knowledge of track topology and the necessity of turnouts that exhibit discriminative features. Depending on this, two scenarios can be distinguished:

- *Side track scenario*: Side tracks are characterized by few stations and turnouts and relatively long spaces of over 1 km between them. The turnouts differ in type and occurrence and comprise individual features like signal box wires, axle counters, and others. Therefore even two turnouts of the same type are distinguishable as seen in Section III-C.
- *Classification yard scenario*: Rail yards and other freight specific surroundings are specified by a significant higher turnout density. They are often laid nearly identically, being of same type. The distances between turnouts are typically small but different. An example setup is shown in Figure 11, that displays only a small area of a classification yard near Ludwigshafen.

Both scenarios are intrinsically solvable with our approach by choosing the HMM input accordingly.

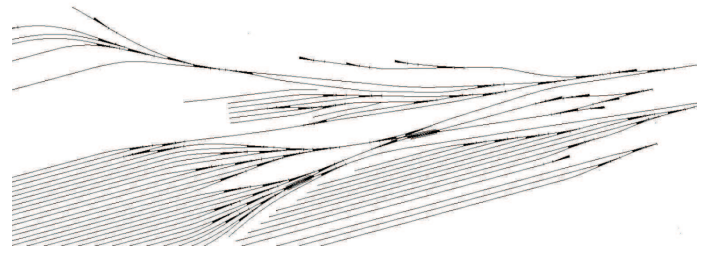


Fig. 11. Cutout of a classification yard near Ludwigshafen showing the vast amount of laid turnouts (black triangles) and its challenges for the shunt yard scenario.

V. EXPERIMENTAL RESULTS

All proposed algorithms were verified with real data. These were recorded on a test train that operates on a secondary track. It incorporates roughly 22km track length, seven stations and an overall of 54 turnouts. In test runs we were passing repeatedly six stations and 23 turnouts in different directions.

A. Velocity Estimation

The verification of the velocity estimation is achieved by comparing the results with a reference measurement system. We use a combination of satellite measurements (GPS) and a dead reckoning inertial navigation system (INS) as reference. Figure 12 shows an example result for a sequence of approx. 300 seconds. It proves that, due to the densely vegetated valley

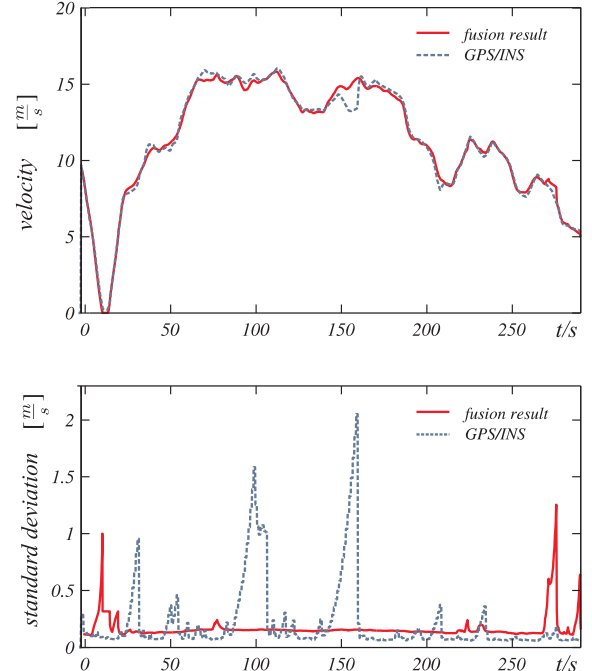


Fig. 12. Comparison of velocity estimation with ECS and reference GPS/INS system.

in which the test track is situated, our approach outperforms the GPS/INS combination in normal or higher velocities. In very low speed regions, filter effects and difficulties in determining the effective correlation length lead to a higher standard deviation for the ECS system. Yet, in an overall

comparison, the ECS approach in Section II-B proves more reliable and robust in average.

B. HMM Results

The algorithms and models described in Sections III-B and III-C are evaluated separately, indicating that a localization is possible with only applying the detection HMM.

1) *HMMs for Detection*: To verify the detection capabilities of our approach, we run 26 test drives (52 passings of the turnouts) on the side track. The employed HMM is constructed according to Figure 6 and consists of 196 states and six submodels as described in Section III-B. The emission densities were estimated out of a set of real turnout signals with manually labeled components. An overall of 845 true turnout events are present on the test track, of which 831 were detected correctly given only one false positive, which yields an overall recognition rate of 98.23% for solely detection. Moreover, we evaluated the driving direction preclassification in the detection step. From the 831 properly detected turnouts 774 were positively classified in the correct driving direction when distinguished in facing and failing. This corresponds to a classification rate of 93.14%.

2) *HMMs for Classification*: The same data set was examined for classification performance. The data set was constrained on turnouts, respectively classes, with at least 20 passings to ensure a correct data set separation in training, validation, and test set. This preselection left an overall of $M = 34$ classes. The output of the second HMM stage yields a specific turnout and therefore node V , as well as a specific driving direction described in Table I, in contrast to the preclassifying of the detection HMM that separates only in facing/failing and an unspecified turnout. The SCHMM emission densities were set with 13 mixture components for a three dimensional input vector consisting of bior wavelet family scales. The average model size N is 232 states, given an average length $\bar{T} = 423.2$ of \mathcal{O} . The maximum iterations of the Baum-Welch algorithm were set to 15, whereas the number of training observation sequences K were set to nine. The overall error is computed on the test and validation sets that contain 277 turnout events. One *false positive* results in a classification performance of 99.64%. Additional examinations showed that the approach is robust up to a velocity estimation error of 15%.

The classification is also nonsensitive against cutting offs at the beginning or end of turnouts, that can occur in the detection HMM. Figure 13 depicts six segments at turnout start and end that were cut off depending on a drawn number from an uniform distribution in multitudes of sleeper distances. It is possible to cut up to 2.5m from the switchblade area and up to 6m from the end, which only reduces the performance to 99.28% (corresponding to two errors out of 277).

C. Localization Results

As outlined in Section IV-B, there are two different application scenarios for rail vehicle localization with SMC.

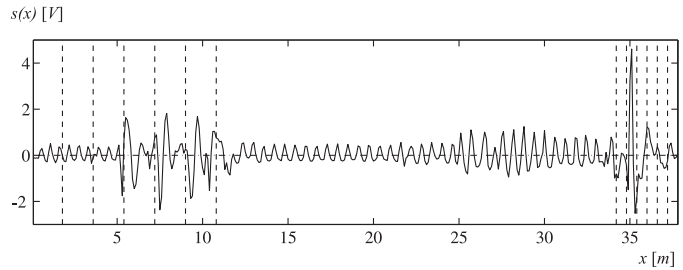


Fig. 13. Bounds of stochastically cut turnout start and end points.

1) *Side Track Scenario*: The side track scenario assumes sufficiently discriminant individual turnouts. Each classified turnout event corresponds to a certain position and the bending in a specific direction. The results in V-B suggest that the probability of the main hypothesis is typically above 95%. This leads to an almost sure convergence of the SMC filter after first classification. The most probable position corresponds to the maximum of $p(\mathbf{x}_t | \mathbf{z}_{1:t})$. This complies to the expectation \hat{x}_t and its variance $\hat{\sigma}$ given with

$$\hat{x}_t = \sum_{p=1}^P w_t^p x_t^p, \quad \text{and} \quad \hat{\sigma} = \sum_{p=1}^P w_t^p (x_t^p - \hat{x}_t)^2, \quad (29)$$

assuming a converged filter and Gaussian errors and influences. A simple threshold is applied to only display the most probable position. The first initialization of the position is achieved in different ways depending on prior knowledge. The initial filter distribution is either set uniform for the whole network or concentrated on only few initial segments of G if rough knowledge is available.

2) *Classification Yard Scenario*: The classification or shunt yard scenario is conditioned on the assumption of turnouts too similar for unambiguous distinction. Still it is possible to detect turnouts and associate the position to geographic coordinates, that can be used as re-calibration for velocity estimation. A Localization based on solely ECS and the detection HMM is possible, whereas the accuracy is based on the capability of the HMM detection to also pre-classify the driving direction in facing and failing sufficiently. The initialization problem is one dimensional within the graph framework and corresponds to the *lost robot problem* described in [11]. Therefore we assume a uniform a priori distribution over the whole considered rail network. Each time a turnout is detected, the distribution converges to only very few or even one hypothesis. This process is visualized in Figure 14 for a simulated station setup nearly identical to a real station from our test track. The HMM detections were simulated at specific time instants. For visualization purposes, the densities are clustered in discrete cells of length 1m and the Z -axis is scaled. The problem to decide whether the train bends right or left is shifted from the classification stage to the distance measurement between turnouts and is therefore done implicitly. On every turnout decision point the particle cloud is separated according to the probability of the bending. This probability can be given a priori in form of a scheduled time table or is assumed uniform if no information is given. Filter convergence is verified with a 5% rejection probability in a

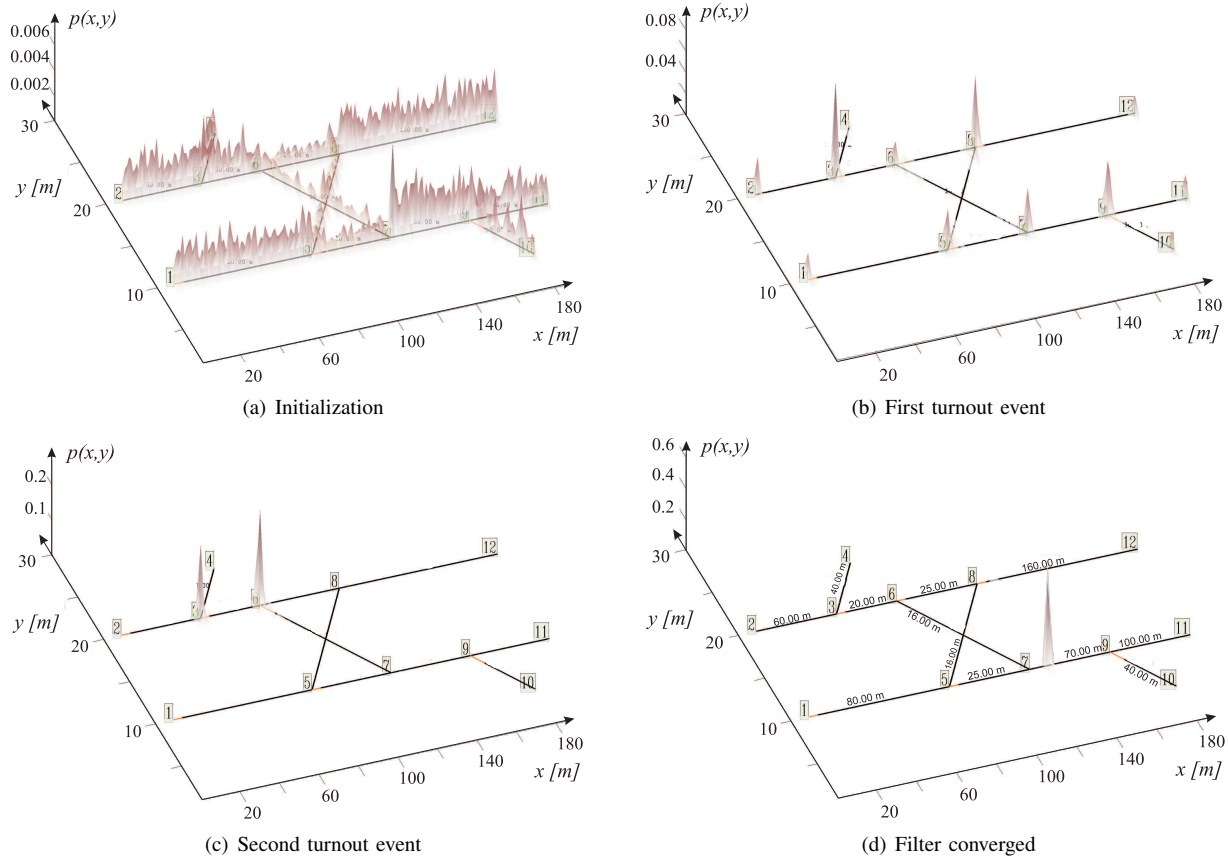


Fig. 14. Example for initialization procedure on an example rail net. The probability $p(x, y)$ is scaled accordingly in each subfigure.

χ^2 test [37]. Each edge E is tested independently on possible positions by applying a *kernel density estimation* (for details see [27]) on the associated particles. The highest maxima of each kernel function are locally separated and tested against the null hypothesis of a Gaussian distribution, relying on the assumptions of Section IV-B. Afterwards, an application dependent threshold for the accumulated probability mass is chosen for the final position decision.

3) *Overall Localization Performance*: The evaluation of the *ECS-only* train tracking with the SMC approach is complicated by the lack of ground truths. In the side track scenario it is obvious that the precision is directly correlated with the results in Section III-C, whereas the tracking in the second scenario depends on a fast convergence. In the side track scenario, estimated velocities and turnout detections were used from real data. The shunt yard scenario V-C1 was validated with the same data sets sparing the classification HMM. Convergence in both scenarios was reached when at least one particle is initialized for each meter of rail and direction. For medium sized shunt yards and stations this corresponds up to 4000 particles. For the whole test track, which represents a common side track and contains 22 km of rail tracks and 7 stations, 44000 particles are sufficient. Results were won by simulating different levels of position errors and its rectification by turnout detection on the one hand, and using test drive data without giving starting position to verify the initialization process on the other. The results were as follows: The filter converges

in every scenario for different configurations. Convergence is assumed after one mode of the whole rail network got 80% of the probability mass. In scenario one this is the case after the first turnout event, scenario two finishes the initialization after an average of 3 turnouts for real data, whereby the results rely on the dissimilarity of the topological segment lengths and the actual position. Tracking after convergence was tested with several wrong initial positions simulating a distance error up to 15m, which corresponds to a maximum distance between two turnouts of 1500m given a velocity error of $\approx 1\%$. These deviations occur after passing inter-station rail tracks and are subsequently reduced by passing turnouts. The algorithm works in real time up to a number of 48000 particles in a Matlab framework given a 2 GHz processor with 2 GB Ram.

VI. CONCLUSIONS

This paper proposes a novel method to localize a train with probabilistic Bayesian methods employing an eddy current sensor and an enhanced topological map. The sensor accurately estimates velocity, based on cross correlation methods, and extracts turnout information. Hidden Markov models are employed for both, detection and classification, one physically motivated, the other trained from data. The subsequent incorporation of continuous velocity estimates and turnout recognition in topological maps is realized with sequential Monte Carlo methods. Our approach copes with a large variety

of turnouts in shape and style, mechanic and electromagnetic disturbances, misclassification and longitudinal drift. The Bayesian tracking is capable to compensate misclassifications and velocity errors to robustly estimate the position and can be used to reconstruct the driven path even in large railroad networks. Our framework can easily be augmented by additional track characteristics or additional sensors if necessary. Future work will incorporate additional ECS information such as the discrete number of sleepers in each segment to combine spatial and time information in one position estimate.

ACKNOWLEDGMENT

The authors would like to thank the German Federal Ministry of Economics and Technology (BMW), Bombardier Sweden, and the Karlsruher Verkehrsbetriebe.

REFERENCES

- [1] P. Winter, J. Braband, and P. de Cicco, *Compendium on ERTMS: European Rail Traffic Management System*. UIC International, 2009.
- [2] F. Böhringer, "Train location based on fusion of satellite and train-borne sensor data," in *Location Services and Navigation Technologies*, vol. 5084, pp. 76–85, 2003.
- [3] P. McIntire and R. C. McMaster, *Nondestructive Testing Handbook, Vol. 4*. The American Society for Nondestructive Testing, Columbus, Ohio, 1986.
- [4] T. Engelberg and F. Mesch, "Eddy current sensor system for non-contact speed and distance measurement of rail vehicles," in *Computers in Railways VII*, pp. 1261–1270, WIT Press, 2000.
- [5] F. Böhringer and A. Geistler, "Comparison between different fusion approaches for train-borne location systems," in *Proc. IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 267–272, 2006.
- [6] F. Kaleli and Y. Akgul, "Vision-based railroad track extraction using dynamic programming," in *Proc. IEEE Conference on Intelligent Transportation Systems*, 2009.
- [7] S. S. Saab, "A map matching approach for train positioning. part i: Development and analysis," in *IEEE Transaction On Vehicular Technology*, vol. 49, pp. 467–475, 2000.
- [8] S. S. Saab, "A map matching approach for train positioning. part ii: Application and experimentation," in *IEEE Transaction On Vehicular Technology*, vol. 49, pp. 476–484, 2000.
- [9] A. Ranganathan, *Probabilistic Topological Maps*. PhD thesis, Georgia Institute of Technology, March 2008.
- [10] B. Kuipers, "The spatial semantic hierarchy," *Artificial Intelligence*, vol. 119, no. 1-2, pp. 191 – 233, 2000.
- [11] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge: The MIT Press, 2005.
- [12] Y. Bar-Shalom, *Multitarget/Multisensor Tracking: Applications and Advances*, vol. 3. Norwood: Artech House, 2000.
- [13] C. Bishop, *Pattern Recognition and Machine Learning*. Information Science and Statistics, 2006.
- [14] F. Puente Lon and T. Engelberg, "Model-based sensing of track components for location of rail vehicles," in *Proc. of the IMTC*, 2005.
- [15] A. Geistler and F. Böhringer, "Robust velocity measurement for railway applications by fusing eddy current sensor signals," in *Proc. IEEE Intelligent Vehicles Symposium*, (New York), pp. 664–669, IEEE, 2004.
- [16] T. Strauss, C. Hasberg, and S. Hensel, "Correlation based velocity estimation during acceleration phases with application in rail vehicles," in *IEEE/SP 15th Workshop on Statistical Signal Processing*, 2009.
- [17] G. Welch and G. Bishop, "An introduction to the kalman filter," in *SIGGRAPH 2001*, p. Course 8, University of North Carolina, 2001.
- [18] W. W. Hay and C. Hay, *Railroad Engineering*. John Wiley & Sons, 1982.
- [19] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Higher Education, 4 th ed., 2002.
- [20] J. Bilmes, "What HMMs can do," tech. rep., University of Washington, 2002.
- [21] R. Durbin, *Biological sequence analysis*. Cambridge University Press, 2002.
- [22] S. P. Chatzis, D. I. Kosmopoulos, and T. A. Varvarigou, "Robust sequential data modeling using an outlier tolerant hidden markov model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, pp. 1657–1669, September 2009.
- [23] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–286, 1989.
- [24] Y. Ephraim and N. Merhav, "Hidden markov processes," *IEEE Transactions on Information Theory*, vol. 48, pp. 1518 – 1569, June 2002.
- [25] M. Ghil, M. R. Allen, A. D. Dettinger, K. Ide, and D. Kondrashov, "Advanced spectral methods for climatic time series," *Reviews of Geophysics*, vol. 40, pp. 1 – 41, 2002.
- [26] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, vol. 5, pp. 260–269, 1967.
- [27] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer, 2001.
- [28] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 2 ed., 1999.
- [29] L. E. Baum, "An inequality and associated maximization technique in statistical estimation for probabilistic functions of markov processes," *Inequalities III*, pp. 1 – 8, 1972.
- [30] X. D. Huang and M. Jack, "Semi-continuous hidden markov models for speech recognition," *Computer Speech and Language*, vol. 3, no. 3, pp. 239 – 251, 1989.
- [31] H. Akaike, "Information measures and model selection," *Bulletin of the international statistical Institute*, vol. 50, pp. 277 – 290, 1983.
- [32] N. Sugiura, "Further analysis of the data by akaike's information criterion and the finite corrections," *Communications in Statistics - Theory and Methods*, vol. A7, pp. 13 – 26, 1978.
- [33] E. T. Jaynes, *Probability Theory the logic of science*. Cambridge University Press, 2003.
- [34] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on communications*, vol. 28, pp. 84–95, January 1980.
- [35] S. Arulampalam, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174 – 188, 2002.
- [36] F. Gustafsson and al., "Particle filters for positioning, navigation and tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 425 – 436, 2002.
- [37] L. Grafakos, *Classical and Modern Fourier Analysis*. Pearson Education, 2004.



Stefan Hensel (S'09) received the Diploma in mechanical engineering at the University of Karlsruhe, Germany, in 2006. His research interest lie in statistical signal processing and pattern recognition with applications in eddy current sensor technology for train application systems. He is currently working at the Institute of Measurement and Control Systems, Karlsruhe Institute of Technology (KIT).



Carsten Hasberg (S'10) studied electrical engineering at the University of Karlsruhe, Germany. He received the Dipl.-Ing. degree at University of Karlsruhe, Germany in 2006. His research interest lies in localization technology for autonomous systems. He is currently working at the Institute of Measurement and Control Systems, Karlsruhe Institute of Technology (KIT).



Christoph Stiller (S'93M'95SM'99) studied electrical engineering at the Universities in Aachen, Germany, and Trondheim, Norway. He received the Dr.-Ing. degree (Ph.D.) with distinction from Aachen University in 1994. He worked in the Research Department, INRS-Telecommunications, Montreal, QC, Canada, and in advanced development for Robert Bosch GmbH, Hildesheim, Germany. In 2001, he became a chaired Professor and Head of the Institute for Measurement and Control Systems at the Karlsruhe Institute of Technology (KIT), Ger-

many. His present interests cover cognition of mobile systems, computer vision, and real-time applications thereof. He is the author or coauthor of more than 100 publications and patents in these fields. Dr. Stiller is Vice President Member Activities of the IEEE Intelligent Transportation Systems Society. He served as Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING (1999–2003) and, since 2004, for the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. In January 2009, he was nominated Editor-in-Chief of the IEEE Intelligent Transportation Systems Magazine. He is the Speaker of the Transregional Collaborative Research Center Cognitive Automobiles of the German Research Foundation..